

EXPLORING AUTOMATION ISSUES IN SUPERVISORY CONTROL OF MULTIPLE UAVS

Heath A. Ruff

Sytronics, Inc., Dayton, OH

Gloria L. Calhoun

Air Force Research Laboratory, Human Effectiveness Directorate, Dayton, OH

Mark H. Draper

Air Force Research Laboratory, Human Effectiveness Directorate, Dayton, OH

John V. Fontejon

Air Force Research Laboratory, Human Effectiveness Directorate, Dayton, OH

Brian J. Guilfoos

Sytronics, Inc., Dayton, OH

ABSTRACT

An evaluation was conducted on a generic UAV operator interface simulation testbed to explore the effects of levels-of-automation (LOAs) and automation reliability on the number of simulated UAVs that could be supervised by a single operator. LOAs included Management-by-Consent (operator consent required) and Management-by-Exception (action automatic unless operator declines). Results indicated that the tasks were manageable, but performance decreased with increased number of UAVs supervised and reduced automation reliability. Performance with the two LOAs varied little and did not show a consistent trend across measures. Analyses indicated that participants typically did not utilize the automation. A follow-on study was conducted that employed shorter LOA time limits. Results showed participants' workload and confidence ratings were less favorable for the shorter limits and they still exercised the automation rarely, although more frequently. Further research is needed to explore the complex relationship between LOAs, time limits, perception of workload, vigilance effects, and confidence.

Keywords: Level of Automation; Supervisory Control; UAV; Reliability; Multi-aircraft control

INTRODUCTION

The majority of present day Unmanned Air Vehicle (UAV) systems require multiple operators to control a single UAV. Reducing the operator-to-vehicle ratio would reduce life-cycle costs and serve as a force multiplier. Thus, automation technology is under rapid development. The envisioned system involves *multiple* semi-autonomous UAVs being controlled by a *single* supervisor. These UAVs will have the capability to make certain higher-order decisions independent of operator input and predefined mission plans. This capability of the UAV 'to decide' constitutes an entirely new tasking on the operator to rapidly judge the appropriateness of

decisions/actions made by the automation and assess their impact on overall mission objectives, priorities, etc. The number of systems to monitor will increase and it will be more of a challenge for operators to maintain situation awareness (SA) through long periods of nominal operations, interjected with short periods of time-sensitive contingency operation.

Unfortunately, it has been documented in studies of manned systems that increasing the use of automation can cause rapid and significant fluctuations in operator workload and can result in loss of operator SA and performance. In fact, there are numerous issues associated with automation management such as task allocation between operator and system, human vigilance decrements, clumsy automation, limited system flexibility, mode awareness, trust/acceptance, failure detection, automation biases, etc. (Parasuraman, Sheridan, & Wickens, 2000). Innovative methods are required to keep the operator 'in the loop' for optimal SA, workload, and decision making. One method that may enhance supervisory control is multiple levels-of-automation (LOAs), whereby each level specifies the degree to which a task is automated. Thus, automation can vary across a continuum of levels, from the lowest level of fully manual performance to the highest level of full automation. Use of higher LOAs might allow for more vehicles to be controlled by a single supervisor. Unfortunately, these high LOAs tend to remove the operator from the task at hand and can lead to poorer performance during automation failures. In contrast, an intermediate LOA that involves both the operator and the automation system in operations may preclude multi-UAV control due to increased operator task requirements. However, it has been hypothesized that an intermediate LOA can improve performance and SA, even as system complexity increases and automation fails. Some research supports this hypothesis (e.g., Ruff, Narayanan, & Draper, 2002) and other results (e.g., Endsley & Kaber, 1999) suggest that there are factors that can impact the benefit of a LOA (e.g., whether task involves option selection versus higher-level cognition). Such results demonstrate the need for more research comparing LOAs in different task environments.

The Air Force Research Laboratory is conducting supervisory control human factors research utilizing a multi-UAV synthetic task environment. The present paper will focus on initial studies examining operator performance and SA with different LOAs and system reliabilities while supervising multiple simulated UAVs.

STUDY ONE

METHOD

Experimental Design

Two LOAs were evaluated. In Management-by-Consent (MBC), the operator had to explicitly agree to suggested actions before they occurred. The automation proposed route re-plans and target identifications, but required operator consent before acting. In Management-by-Exception (MBE), the system automatically implemented suggested actions after a preset time period, unless the operator objected. The settings for the MBE LOA (time limit until override) and the low/high reliability levels were: image prosecutions: 40 sec, 75/98%; route re-plans: 15 sec, 75/100%. The experiment employed a mixed design: 1 between-subjects variable (automation reliability, low/high) and 3 within-subjects variables (number of UAVs, LOA, and monitor arrangement). (Monitor arrangement will not be addressed here due to space restrictions.) The

LOA variable was blocked and counterbalanced. UAV number (2 or 4) and monitor arrangement (horizontal or vertical) were also counterbalanced. After completion of training on the displays and all tasks/variables, each of the 16 participants completed 8 experimental trials, one sixteen-minute trial with each combination of independent variables.

Multi-UAV Synthetic Task Environment

The MIIRO (Multi-Modal Immersive Intelligent Interface for Remote Operation; Tso, et al., 2003) testbed was utilized, consisting of two monitors, a keyboard, and mouse (Figure 1). One monitor (Figure 2, left) presented the Tactical Situation Display showing the color coded UAV routes, suggested route re-plans, waypoints, targets, threat rings, and any unidentified aircraft. As each UAV passed a target, its camera took images and these appeared in the queue at the bottom of the Image Management display (Figure 2, right). The image in the top row of the queue was displayed. Suspected hostile targets within the image were highlighted by the automatic target recognizer (ATR) with red squares.



Figure 1. Multi-UAV Task Environment.

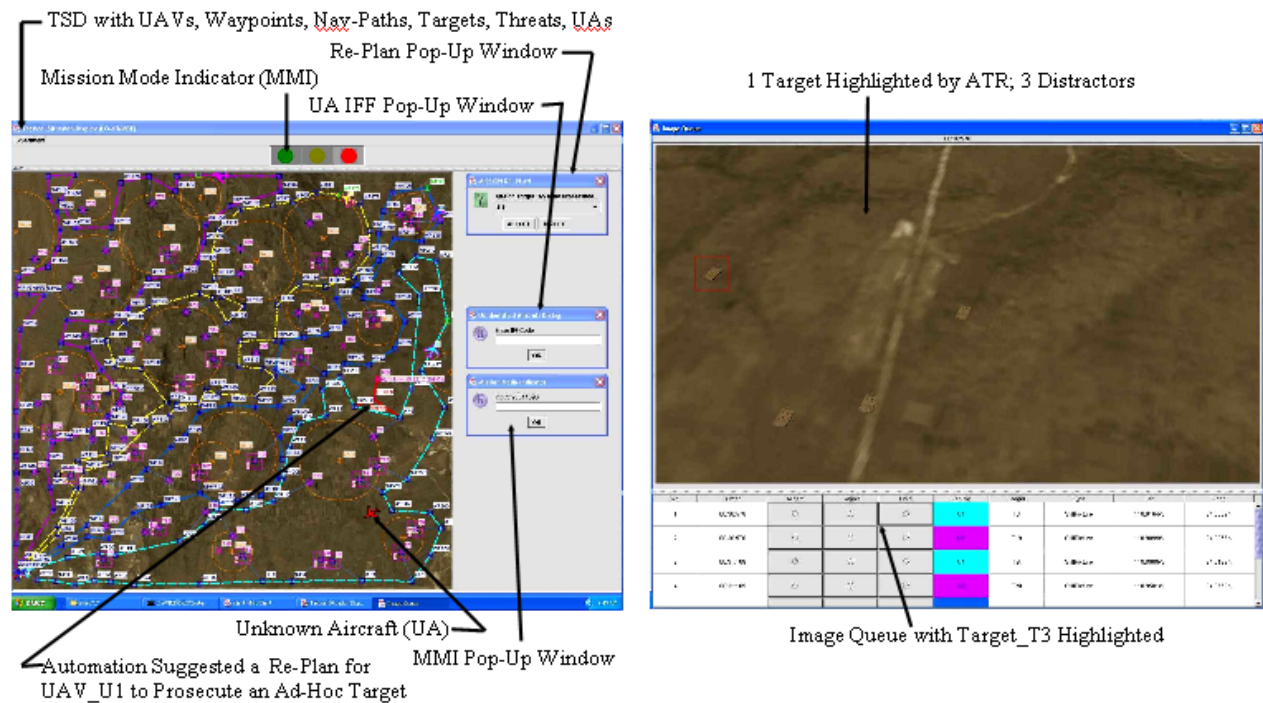


Figure 2. Examples: Tactical Situation Display (left) and Image Management Display (right).

Mission/Operator Tasks

Participants were required to respond to several types of events, listed in order of priority:

- Unidentified Aircraft (2 per mission). This task emulated having a highly unexpected, non-routine, high-priority event occur during a mission. When participants saw a red airplane icon appear, the response was to click on the symbol and enter a code in a pop-up window.
- Route Re-Plans (16 per mission). When alternate routes were suggested by the automation in response to ad-hoc targets and threats, participants were required to inspect the alternate route and make a decision to accept or reject the re-plan in a pop-up window, based on whether the re-plan crossed another threat or another UAV's route.
- Image Prosecutions (per mission: 34 (2 UAVs), 66 (4 UAVs)). Participants were required to view the image in the top window and verify that red boxes were only around targets (versus distractors). Participants could add or delete boxes by clicking on the items, if there were errors. Then participants made an accept/reject decision by clicking the appropriate box.
- Mission Mode Indicator (MMI) (per mission: 16 (2 UAVs), 32 (4 UAVs)). This secondary monitoring task was used to represent the various contingency management panels that will likely exist in future stations. The panel's green light meant everything was operating normally. When this light extinguished and either the yellow or red light activated, then participants' response was to click on the panel and make an entry in a pop-up window.

RESULTS/DISCUSSION

Data recorded included time and accuracy in responses to: 1) image prosecutions, 2) proposed re-plans, and 3) system state changes and unknown aircraft. Workload, SA, and trust ratings were also collected. Results indicated that the tasks were manageable, but performance and subjective ratings decreased with:

- Increased number of UAVs: For image prosecutions, route re-plans, and MMI tasks, participants' average completion times were faster with 2 UAVs than 4 (all $p < .01$) and less time was spent in threat zones ($p < .05$). With the 2 UAV condition, participants were also more likely to respond before the automation acted ($p < .01$). The subjective ratings indicated that participants viewed the 4 UAV condition as higher workload, more difficult, and less trustworthy (all $p < .01$).
- Reduced automation reliability: Fewer images were prosecuted and more errors were made ($p < .01$) in the Low Reliability level compared to the High level. The subjective data also indicated that the participants had less trust when Reliability was Low ($p < .05$).

Performance between the two LOAs varied little and did not show a consistent trend across measures. The design dictated that trials with the MBC automation never timed-out. With MBE, participants typically responded rather than let the action automatically occur. In fact, image prosecution time averaged 12 seconds for both LOAs, much shorter than the criterion time limits employed. Thus, the results pertaining to LOAs are questionable, as the automation was not utilized as designed. Rather, the results suggest that the time criteria employed in the LOAs should be shortened significantly, to determine whether automation is a benefit in this simulated task environment. A follow on study was conducted to evaluate this change.

STUDY TWO

METHOD

Two of the three variables were the same as the first study: Automation Reliability (low/high; between-subjects) and LOA (MBC/MBE; within-subjects). A third (within-subjects) variable was Time Limit for the LOA (“short/long”). The short/long time limits to override were: image prosecutions (15/40 sec) and route re-plans (10/15 sec). The LOA variable was blocked and the order counterbalanced across subjects. The order of the Time Limit levels was counterbalanced within each LOA block. For all trials, there were 4 UAVs and the monitors were arranged horizontally. After training, each of 16 participants completed 4 experimental trials, one sixteen-minute trial with each combination of the independent variables.

All other procedures were the same as that used in the first study, except for how the route re-plan task was implemented. In Study One, participants were only required to inspect whether the re-plan crossed the path of another UAV or a threat zone. To better simulate the cognitive effort anticipated in operational missions, Study Two’s re-route task required participants to view three readouts in a pop-up window that gave two fuel levels and the UAV’s “resources” (low/medium/high). The accept/reject criteria was based on a mathematical relationship between these variables (e.g., if Fuel A plus .5 Fuel B is greater than 5 *and* Resources = Low, then Re-route should be accepted).

RESULTS/DISCUSSION

The efficacy and flexibility of the testbed were demonstrated by the successful change in the route re-plan task. (Average completion time, with longer Time Limit, was longer in Study Two (by 2.2 sec), presumably reflecting the changes in this task to increase its cognitive difficulty.) Also different in Study Two, only one measure showed a significant effect of Reliability: the percentage of images correctly prosecuted was less for Low, compared to High ($p < .01$). In regards to LOA, there were no significant differences in the performance and subjective measures, except as a function of the Time Limit variable. Participants’ difficulty and workload ratings were similar for the two Time Limits for MBC LOA. With MBE, however, their ratings indicated the shorter limit was higher workload (Figure 3, left) and more difficult (both measures, $p < .05$). The participants’ ratings may reflect the fact that their average time to complete image prosecutions was faster with the shorter time limit in MBE ($F(1,14) = 5.256, p < .05$; Figure 3, right) than the other three combinations of LOA and Time Limit. These findings may be related to the participants’ ratings of less confidence with the shorter time limits ($p < .01$) and the nature of the LOA. In MBE, if the participant didn’t respond to images before the time limit, they were automatically prosecuted. The fact that an erroneous action could occur, and more likely with the shorter time limit, may have pressured participants to respond faster and view it as higher workload. Thus, although MBE was hypothesized to be a workload reducer, it actually appeared to add to perceived workload.

Time Limit was also key in terms of the frequency in which the automation was exercised. Both image prosecution and route re-plans were more likely to activate automatically in trials with the shorter limit (e.g., 12.4% of the image prosecutions were automated in trials with the shorter limit, 1% with longer limit, the latter similar to Study One that employed a

similar time limit). Yet, most re-plans and image prosecution tasks were completed manually, in less time (7.2 and 11.7 sec, respectively) than the available Shorter Time limits (10/15 sec).

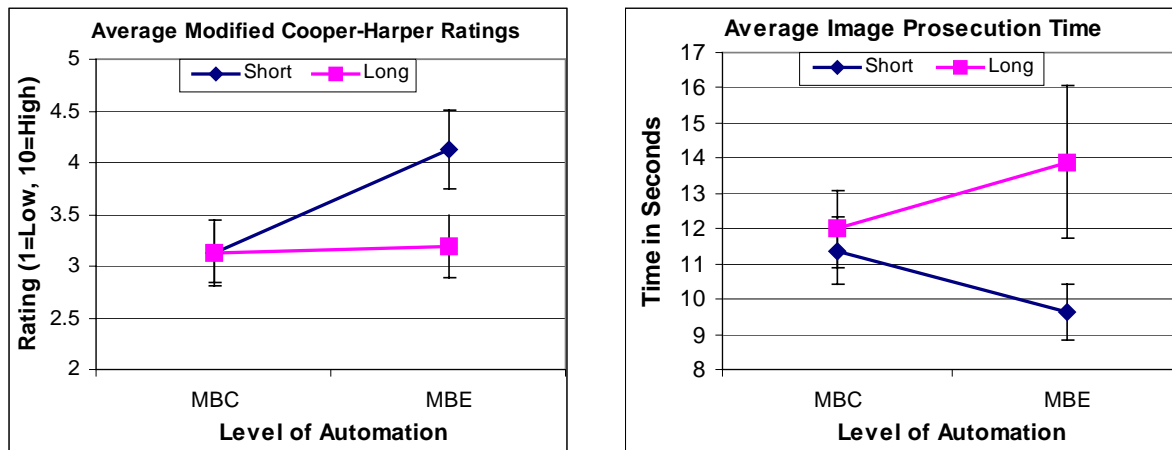


Figure 3. For each LOA (Management-by-Consent and Management-by-Exception) and Time Limit (Short/Long): Average Modified Cooper-Harper Rating for Workload (left) and Average Image Prosecution Time (right) with Standard Error of the Mean.

CONCLUSIONS

The rarity of automated actions, together with the increased workload and decreased re-plan and image prosecution times and confidence ratings with the shorter time limit, suggests that the participants preferred to respond manually rather than rely on the automation. At the very least, these results illustrate the complex relationship between LOA, time limits, and perception of difficulty and confidence. Moreover, participants' inclination to exercise the automation may increase in longer trials where vigilance effects are more likely to occur. Further research is needed before an optimal operator system design can be determined for supervision of multi-UAVs. This research will also explore the utility of additional LOAs that are: 1) contingency/task specific and 2) changeable during a mission, to better explore the utility of context-sensitive automation and decision aiding in UAV supervisory control.

REFERENCES

- Endsley, M. R. and Kaber, D. B. (1999). Level of automation effects on performance, situation awareness and workload in a dynamic control task. *Ergonomics*, 42(3), 462-492.
- Parasuraman, R., Sheridan, T. B., and Wickens, C.D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans*, 30(3), 286-297.
- Ruff, H. A., Narayanan, S., and Draper, M. H. (2002). Human interaction with levels of automation and decision-aid fidelity in the supervisory control of multiple simulated unmanned air vehicles. *Presence: Teleoperators and Virtual Environments*, 11(4), 335-351.
- Tso, K. S., Tharp, G. K., Tai, A. T., Draper, M. H., Calhoun, G. L., & Ruff, H. A. (2003, October). A human factors testbed for command and control of unmanned air vehicles. *Digital Avionics Systems 22th Conference*, 8.C.1-1-12.